

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-320400

(43) 公開日 平成10年(1998)12月4日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/40

3 7 0 D

審査請求 未請求 請求項の数12 O L (全 16 頁)

(21) 出願番号

特願平9-126704

(22) 出願日

平成9年(1997)5月16日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 長坂 晃朗

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 宮武 孝文

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 弁理士 小川 勝男

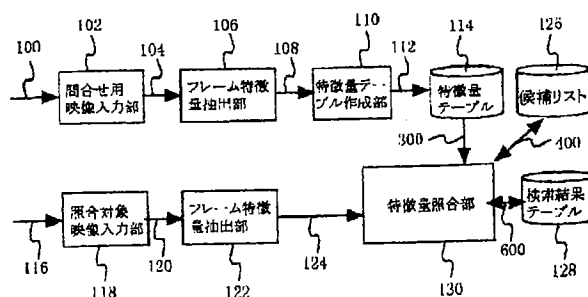
(54) 【発明の名称】 映像の検索方法および装置

(57) 【要約】 (修正有)

【課題】 放送中のビデオ映像またはデータベース中のビデオ映像を高速に検索する方法、ならびに、映像中の部分映像の同一性に基づき分類整理する映像の自己組織化方法を提供する。

【解決手段】 映像を逐次フレーム毎に入力する手段102と、入力したフレーム画像から特徴量を抽出する手段106と、該特徴量を時間軸方向に圧縮する手段と、該圧縮された特徴量を逐次蓄積する手段114と、該蓄積手段内の特徴量と別途入力したフレーム画像の特徴量との間で逐次照合する手段130と、暫定的に照合の途中経過を記憶しておく手段を有し、該照合手段では照合の途中経過を更新しながら、一致する映像部分を探す手段とを有する。さらに、一致する映像部分が複数存在した場合には、それらを組にして対応づけて記憶格納する手段128とを有する。

図 2



## 【特許請求の範囲】

【請求項1】映像を逐次フレーム毎に入力する手段と、入力したフレーム画像から特徴量を抽出する手段と、該特徴量を時間軸方向に圧縮する手段と、該圧縮された特徴量を逐次蓄積する手段と、該蓄積手段内の特徴量と別途入力したフレーム画像の特徴量との間で逐次照合する手段と、暫定的に照合の途中経過を記憶しておく手段を有し、該照合手段では照合の途中経過を更新しながら、一致する映像シーンを探そうにしたことを特徴とする、映像検索装置。

【請求項2】映像を逐次フレーム毎に入力する手段と、入力したフレーム画像から特徴量を抽出する手段と、該特徴量を時間軸方向に圧縮する手段と、該圧縮された特徴量を逐次蓄積する手段と、該蓄積手段内の特徴量と該入力フレーム画像の特徴量との間で逐次照合する手段と、暫定的に照合の途中経過を記憶しておく手段を有し、該照合手段では照合の途中経過を参照しながら、一致する映像シーンを探そうにしたことを特徴とする、映像検索装置。

【請求項3】請求項1または2記載の特徴量はフレーム画像の輝度または色の統計量を利用したことを特徴とする映像検索装置。

【請求項4】請求項1または2記載の特徴量の時間軸方向の圧縮は、最初に抽出した特徴量とそれに引き続くフレーム画像の特徴量が許容範囲内となる時間区間毎に区切ったことを特徴とする映像検索装置。

【請求項5】請求項1または2記載の照合は、特徴列同士的一致が一定の時間区間以上続くと判定されたとき、照合がとれたとすることを特徴とする映像検索装置。

【請求項6】請求項1または2記載の照合の途中経過記憶は、照合の可能性のある先頭フレームの番号を仮に登録し、照合が続く場合はその時間を更新し、照合の可能性のなくなった時は、該当するフレーム番号を削除するようにしたことを特徴とする映像検索装置。

【請求項7】請求項1または2記載の照合において、入力された特徴列に対して、複数箇所、同一として照合された映像部分は、それらを組にして対応づけて結果を記録するようにしたことを特徴とする映像検索装置。

【請求項8】請求項2記載の映像検索方法および装置において、複数箇所、同一として照合された映像部分をもとに、放送中の連続番組を分類するようにしたことを特徴とする映像検索装置。

【請求項9】請求項2記載の映像検索方法および装置において、複数箇所、同一として照合された映像部分とその時間長をもとに、放送中の特定の映像、特にCMを検出するようにしたことを特徴とする映像検索装置。

【請求項10】ディスプレイと、外部から画像の入力制御と該ディスプレイへの画像の表示制御を行う制御手段と、メモリとを有する情報処理装置において読みとり可能な記録媒体であって、該制御手段において、

フレーム単位で動画像を入力してメモリに蓄積する処理、

新たにフレームが入力された場合に、該フレームが最新フレームの所定時間長の動画像区間1と同一の動画像区間2を、上記メモリから検索して上記動画像区間1及び2の関連づけを行う処理、

入力された動画像をディスプレイで簡易表示する場合に、関連づけられた上記動画像区間1及び2を他の動画像区間とは区別して表示する処理、を行わせるコンピュータプログラムを記憶した記憶媒体。

【請求項11】ディスプレイと、

処理プログラムとデータ保持領域を有するメモリと、

上記処理プログラムに従って、画像の入力処理、画像の検索処理、該ディスプレイへの検索画像の表示処理、を行う制御手段と、

を有し、

上記画像の入力処理では、入力された動画像がフレーム単位でメモリに蓄積され、

上記画像の検索処理では、新たにフレームが入力された場合に、該フレームが最新フレームの所定時間長の動画像区間1と同一の動画像区間2を、上記メモリから検索して上記動画像区間1及び2の関連づけが行われ、

上記ディスプレイへの検索画像の表示処理では、入力された動画像をディスプレイで簡易表示する場合に、関連づけられた上記動画像区間1及び2が他の動画像区間とは区別して表示される、ことを特徴とする情報処理装置。

【請求項12】画像を入力するカメラと、

該画像の入力処理部と、

該カメラから入力された画像を記憶する記憶部と、

該記憶部に記憶された画像を再生出力する出力処理部と、

該画像を表示する表示部とを有する画像録画装置において、

入力された画像の特徴量をフレーム毎に抽出する特徴抽出部と、

該抽出された特徴をテーブル化して保持するメモリ領域と、

入力された画像の特徴量と上記テーブル上の特徴量とを照合する処理と、所定の照合条件に合致する特徴量を有するフレーム同士を類似の画像として関連づける処理を制御する制御部と、を有することを特徴とするビデオカメラ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は放送中のビデオ映像またはデータベース中のビデオ映像の検索方法に関し、特にビデオ映像の特徴を手掛かりとした検索が高速に行えるビデオ映像の検索方法に関する。

【0002】

【従来の技術】近年のマルチメディア情報処理システムではビデオやテキスト等様々な情報を蓄積してユーザに提示することが可能であるが、それらを検索する場合、キーワード等の言語による検索方法が主流となっている。この場合はキーワード付けの作業が必要であり、大量の情報量を持つビデオ映像の1枚1枚のフレームについてキーワード付けを行うことは非常にコストが大きい。さらにキーワードはデータベース構築者が自由に付けられるものであるから、ユーザの視点がデータベース構築者と異なる場合にはこれらのキーワードが役に立たなくなるという問題がある。このような事情から、キーワード以外に画像独自の特徴から検索を行いたいという要求があるが、画像の特徴量を手掛かりとして検索するためには膨大なフレーム数からなるビデオの特徴量ととの問い合わせ用ビデオの特徴量との高速照合技術が必要になる。ビデオ映像に対して適用可能なだけの高速な照合技術としては、これまで特開平7-114567号「ビデオの検索方法および装置」が提案されているが、この方法は、すべてのフレームについて照合するのではなく、映像中のカットの変わり目の画像についてのみ照合することで処理量の低減を図っている。これによって、放送中の映像に対しても照合を行えるだけの高速性を実現しているが、その反面、1つのカットのみで構成されるシーンや、編集等によってカットの変わり目が前後に変動したシーンに対して、うまく照合ができないという問題点があった。また、検索の際には、他の一般的なデータベースシステムと同様、検索キーとして指定されたシーン以外は探さないため、シーンの検索が必要になるたびに、膨大な量の映像情報についてその最初から最後まで繰り返し照合を行う必要がある。シーンの照合処理には、特徴量の抽出や読み出しの処理をはじめ、検索するシーンが違って共通に行われる処理が多々あり、そのような処理を重複して行うことにはまだ無駄がある。

【0003】

【発明が解決しようとする課題】本発明が解決しようとする第一の課題は、映像検索のためのキーワード付け作業を行うことなく、検索対象映像の特徴量と、問い合わせのために用意するサンプル映像の特徴量とを高速に照合し、フレーム精度で同一の区間を検出できる映像検索方法を提供することにある。また対象映像は放送中およびデータベース中のいずれの映像も対応可能とする。

【0004】本発明が解決しようとする第二の課題は、対象映像中に存在する同一シーンを、それが検索キーとして予め指定されているか否かにかかわらず、一様に対象映像の入力と同時に検出しておく手法を提供することにある。

【0005】本発明の第三の課題は、映像の撮影中に時々刻々と入力されている画像列を録画する際に、録画済みの画像と照合して、一致する画像と関連づけながら録画するビデオカメラを提供することにある。

【0006】

【課題を解決するための手段】上記課題を解決するために、本発明では、照合したいビデオ映像について、各フレームの特徴量が互いに特定の範囲内の変動幅で収まる区間単位に分割し、該区間ごとに1つもしくは複数の特徴量を抽出し、該区間の該映像中における位置を表わすアドレス情報と対応づけて記憶格納しておき、その上で、検索対象であるビデオ映像から順次1枚ずつフレーム画像を入力し、該フレーム画像の特徴量を順に並べた任意の時点の特徴列と、該記憶格納された映像を構成する各区間の特徴量を区間の長さ分ずつ順に並べた特徴列とが、互いに同等と判定できる特定の長さ以上の部分を持つ場合に、その部分を同一の映像として検出する。このとき、区間の先頭から同等の場合は、該区間に対応するアドレス情報を、また、区間の途中から同等と判定された場合には、その区間の先頭からの相対位置を求め、該区間に対応するアドレス情報を補正した値を検索結果として出力する。また、検索対象として入力されたフレーム画像列について、フレームの特徴量が互いに特定の範囲内の変動幅で収まる区間ごとにまとめ、該区間ごとに1つもしくは複数の特徴量を抽出し、該区間の該対象映像中における位置を表わすアドレス情報と対応づけた情報も記憶格納し、次回からの照合対象映像の中に追加する。さらに、入力されたある特徴列について、同一として検出された映像部分が複数存在した場合には、それらを組にして対応づけて記憶格納する。

【0007】以上の検索方法を実現する装置は、任意のビデオ映像に対して、フレームの特徴量が互いに特定の範囲内の変動幅で収まる区間単位に分割する手段と、該区間ごとに1つもしくは複数の特徴量を抽出する手段と、該区間の該映像中における位置を表わすアドレス情報と対応づけて記憶格納する手段と、検索対象であるビデオ映像から順次1枚ずつフレーム画像を入力する手段と、該フレーム画像の特徴量を順に並べた任意の時点の特徴列を保持する手段と、該記憶格納された映像を構成する各区間の特徴量を区間の長さ分ずつ順に並べた特徴列を生成する手段と、それらの特徴列が互いに同等と判定できる特定の長さ以上の部分を持つかどうかを判定する手段とから構成される。また、区間の先頭から同等と判定された場合は、該区間に対応するアドレス情報を、また、区間の途中から同等と判定された場合には、その区間の先頭からの相対位置を求め、該区間に対応するアドレス情報を補正した値を検索結果として出力する手段も併せ持つ。また、検索対象として入力されたフレーム画像列について、フレームの特徴量が互いに特定の範囲内の変動幅で収まる区間ごとにまとめる手段と、該区間ごとに1つもしくは複数の特徴量を抽出する手段と、該区間の該対象映像中における位置を表わすアドレス情報と対応づけた情報も記憶格納し、次回からの照合対象映像の中に追加する手段とを有する。さらに、入力された

ある特徴列について、同一として検出されたシーンが複数存在した場合には、それらを組にして対応づけて記憶格納する手段も有する。

【0008】

【発明の実施の形態】以下、本発明の一実施例を図を用いて説明する。

【0009】図1は、本発明を実現するためのシステム構成の概略ブロック図の一例である。1はCRT等のディスプレイ装置であり、コンピュータ2の出力画面を表示する。コンピュータの出力が音声である場合には、13のスピーカを通して出力する。コンピュータ2に対する命令は、ポインティングデバイス3やキーボード4を使って行うことができる。5のビデオ再生装置は、光ディスクやビデオデッキ等である。ビデオ再生装置5から出力される映像信号は、逐次、6のビデオ入力装置によってデジタル画像データに変換され、コンピュータに送られる。また場合によっては放送中の映像を取り込むことも可能であり、放送受信機7からの映像信号をビデオ入力装置6へ入力する。5のビデオ再生装置の代わりに、デジタルデータとして映像を記録するビデオサーバやデジタルビデオ等を用いる場合には、6のビデオ入力装置は不要か、もしくは圧縮記録された画像データを伸張して非圧縮画像データに変換する機能を司る。放送受信機7についても、放送がデジタル方式の場合には同様である。コンピュータ内部では、デジタル画像データは、インタフェース8を介してメモリ9に入り、メモリ9に格納されたプログラムに従って、CPU10によって処理される。10が扱うビデオがビデオ再生装置5から送られている場合は各フレーム画像には、ビデオの先頭から順に番号（フレーム番号）が付けられている。フレーム番号を制御線11によってビデオ再生装置に送ること  
30  
で、当該場面のビデオが再生するよう制御可能である。また放送受信機7から送られてくるビデオの場合、フレーム番号はないので必要に応じて、処理開始時を0とした連番や時刻等を記録してフレーム番号の代わりに使用する。コンピュータ内部処理の必要に応じて、各種情報を12の外部情報記憶装置に蓄積することができる。メモリ9には、以下に説明する処理によって作成される各種のデータが格納され、必要に応じて参照される。

【0010】図2は、本発明の映像検索処理の処理概要を示す全体ブロック図である。この処理は、コンピュータ2の内部で実行される。処理プログラムはメモリ9に格納され、CPU10において実行される。以下では、各部がCPU10によって実行されるソフトウェア手順として記述されることを前提に説明を行うが、本手順と等価な機能をハードウェアによって実現できることは言うまでもない。なお、以下の説明において、ソフトウェアにより行われる処理が便宜上ブロック化して示した。従って、例えば、図2において問い合わせ用映像入力部は問い合わせ用映像入力処理を示している。本実施例で  
40  
50

は、見つけ出したいシーンの映像（以下、問合せ映像と呼ぶ）100は、検索に先立って予め、問合せ用映像入力部102によってフレーム毎に逐次入力され、メモリ9に一時的に格納される。フレーム特徴量抽出部106は、メモリ9のフレーム画像104から特徴量108を抽出する。特徴量テーブル作成部110は、特徴量が許容変動範囲内にある一続きの区間ごとに、その特徴量と、その先頭フレーム番号とを対にして、特徴量テーブル112を作成し、記憶装置114に記録する。検索対象であるビデオ映像116も、問合せ映像と同様にして照合対象映像入力部118によってフレーム毎に逐次入力され、メモリ9に一時的に格納される。フレーム特徴量抽出部122は、メモリ9のフレーム画像120から特徴量124を抽出する。ここで、122は106と全く同じ処理を行う。特徴量照合部130は、122から次々と送られてくる特徴量124の最新の時系列の並びと、記録された特徴量テーブル300（データ内容は112と同一）とが一致するか比較照合する。照合の途中経過は、記憶装置126に後述する候補リスト400の形で記憶され、新しいフレームの入力ごとに更新される。もし、特徴量が一致すれば、その特徴量テーブルに対応する映像区間を、後述する検索結果テーブル600として記憶装置128もしくはその他の処理装置に出力する。このとき、その検索された映像に何らかの名称や属性等が対応づけられていた場合、その名称や属性を出力することも当然可能である。

【0011】次に、上記各部で行われる処理について、より詳細に説明する。

【0012】図3は、問合せ用映像が入力されて特徴量テーブルが作成されるまでの一連の流れ（100～114）を示したものである。この処理の目的は、より多くの種類の問合せ用映像を記憶しリアルタイムで一度に照合できるように、問合せ用映像を、その特徴を表現する最低限度の情報に圧縮することにある。具体的には、まず次々と入力されるフレーム画像から特徴量を抽出する。このとき、特徴量は、フレーム画像全体の平均色など数バイト程度で表現できる情報に限定する。さらに、得られた特徴量の時系列の並びを、許容変動範囲内にある一続きの区間ごとにまとめ、それぞれの区間につき1つの特徴量で代表させる。図中、A'、A''と表記したのは、Aを基準として、その特徴量の値からの差の絶対値が特定閾値以内の特徴量であることを示している。入力された映像の各フレームには、 $t_1$ から順に $t_2$ 、 $t_3$ 、…とフレーム番号が添えられており、各区間の先頭フレームのフレーム番号 $t_1$ 、 $t_2$ 、 $t_3$ 、…と特徴量A、B、C、…とを対にしてリストを作り、特徴量テーブルとする。ここで、映像は1秒あたり30枚ものフレーム画像から構成されるため、探したい映像の種類にも依存するが、平均区間長が10フレームとして、わずか数秒分のシーンからでも10個以上の特徴量からなる順列パターンが

得られる。これに各区間の長さも制約条件に加えれば、このときの特徴量テーブルの順列組合せの数はきわめて大きな数になり、多数の映像中にあっても十分に1つのシーンを特定できるだけの性能が期待できる。

【0013】図4は、検索対象のビデオ映像と、予め記憶してある問合せ用映像との間の照合の様子(特徴照合処理130)を模式的に示したものである。上述のように検索対象映像は、逐次フレーム画像入力され、特徴量が抽出される(116から124)。一方、特徴量テーブルの形で圧縮されていた問合せ用映像は、照合時(130)には、各区間の長さ分ずつ特徴量がならべられ、ラン単位からフレーム単位の特徴列に戻される。照合にあたっては、対象映像から入力されたばかりの最新のフレームを最末尾とする特定閾値以上の長さの特徴列と一致する特徴列を持つ問合せ用映像を検索結果として返す。このとき、完全一致だけでなく部分的な特徴列の一致も検出し、その一致した部分の長さが同閾値以上の場合には、それも検索結果として返すようにする。これにより、編集等で長さが微妙に異なっているシーンに対しても正しく検索ができるようになる。

【0014】図5は、本発明における照合処理をより詳細に示したものである。上述したような不定長の特徴列の照合を行う場合、単純に行うと、対象映像からフレーム画像が新しく入力される都度、図6に示したような、様々なフレーム長を想定した比較照合を繰り返さなければならない。このときのフレーム間比較の数は、図から明らかなように極めて膨大であり、とりわけ1/30秒に一度の割合で続々と新しいフレームが入力されるようなリアルタイム照合用途には不向きである。これは、フレーム入力のたびごとに以前の照合処理とは全く独立に照合処理を行っているためであり、直前の処理で、ある程度の長さ分の一致が確認されたとしても、その情報を次の照合処理に活かすことができない。そこで本発明では、1回のフレーム入力で行う照合処理を削減し、フレーム入力の都度、直前までの処理を補足するように照合処理を段階的に行うアプローチをとる。具体的には、次のようにして照合を行う。

【0015】(1)あるフレームが対象映像から入力されたとき、そのフレームの特徴量と同じ特徴量が問合せ用映像中にあるかどうかを探して、見つかったフレームをすべて候補として一時的に記憶する。

【0016】(2)次のフレームが対象映像から入力されたとき、そのフレームの特徴量が、直前に候補として記憶したフレームの直後のフレームの特徴量と一致するかどうかを調べる。

【0017】(3)一致しなかった場合には候補から外し、また、今回入力されたフレームの特徴量と同じ特徴量を持つフレームを新たに候補として追加する。このとき、候補から外されたフレームが、それまでに特定閾値以上の長さ分だけ一致し続けていた場合には、そのフレ

ームを先頭にした一致区間を検索結果として出力する。

【0018】(4)以上を繰り返す。

【0019】以下、図5の例に従い、本発明による照合の原理を具体的に説明する。

【0020】まず対象映像から新規のフレーム入力があり、特徴量Xが得られたフレーム(1)について考える。問合せ用映像の中に特徴量Xは存在しないので、何もしない。フレーム(2)も同様である。フレーム(3)が入力されて特徴量A'が得られると、A'と一致する特徴量Aが問合せ用映像中に存在するので、特徴量Aを持つ問合せ用映像中のすべてのフレーム①～④を候補に入れる。対象映像から今後入力されるフレームの特徴量の現れかたによっては、これらの候補フレームのいずれも、そのフレームを先頭とした一続きの区間が検索シーンとなる可能性を秘めている。図5の下の方において、フレーム(3)の行に書き込まれている①～④が、この時点で候補に入れられた問合せ用映像中のフレームを示している。続くフレーム(4)でも特徴量A'が得られるが、まず前回候補に入ったすべてのフレームについて、その次のフレームが特徴量が一致するかどうかを調べる。結果は、①～③のフレームについては一致するが、④のフレームについては、次のフレーム⑤で特徴量がBに変わっているので一致しない。表中4行目に×と表記されている部分がそれで、フレーム(3)で候補に入った④については、この時点で候補から外される。また同時に、フレーム(4)における候補として、表中4行目に、(3)と同じ①～④が新たに加えられる。(3)行目で加えられた①～④と(4)行目で加えられた①～④とは、同じフレームではあるが、照合の候補としては全く別物として扱う。さらに、フレーム(5)ではB''が得られ、ここで(3)で候補に入った①②と、(4)で候補に入った①～③が候補が外される。そして、同様にして⑤と⑥が、この時点で候補に入れられる。以上の処理を、対象映像からフレームを入力するたびに繰り返すと、フレーム(8)の段階まで一致し続けている候補は、(3)で候補に入った③と、(4)で候補に入った④、(5)で候補に入った⑤、(6)で候補に入った⑥、(7)で候補に入った⑦だけである。そして、フレーム(9)が来て照合がとれなくなった時点で、対象映像のフレーム(3)～(8)と、問合せ用映像の③～⑧とが最も長く一致する区間であったことがわかる。これらの結果は、先に図6で示した従来方法を用い、フレーム(8)を基点として順に長さを変えながらシーンの照合を調べた場合の照合結果と一致する。この場合、問合せ用映像のフレーム数をnとすると、1回のフレーム入力あたりに行わなければならないフレーム間比較の回数は、図6に示したように $n(n+1)(n+2)/6$ 回となり、その計算量のオーダーは $O(n^3)$ となる。しかし、本手法によれば、(1)新たに入力されたフレームの特徴量と、候補フレームの次フレームの特徴量との一致を調べる回数cと、(2)問合せ用映像の中に、新たに入力されたフレームの特徴量と

同じ特徴量があるかを調べる回数 $n$ 、の和だけでよく、一般に $n \gg c$ となってオーダーは $O(n)$ となる。この差は、直前のフレームまでの処理結果をもとにして、現在のフレームを加えた結果を求める帰納的な手法としたことによる。また、 $n$ は、前述した特徴量テーブルの利用によって、もともとのフレーム数よりも少なくでき、さらに高速な照合が期待できる。そしてまた、この検索結果は、明らかにフレーム精度で位置決めができています。

【0021】以上の説明では、問合せ用映像が一つの場合を想定して説明したが、複数の場合でも問題なく適用できる。フレーム入力ごとの照合にあたって、問合せ用映像の数だけ上記の処理を繰り返せばよいだけである。但し、図7に示したように、問合せ用映像のそれぞれに同じ映像部分が含まれていながら、それらが編集のしかたが異なるなどの影響で微妙に前後が違っている場合がある。図では、①、②、③の3種類が示されている。1つの問合せ用映像の中に複数の同一映像部分が含まれている場合も同様である。問合せ用映像の中に一致する部分があるかどうかだけが必要となきには問題にならないが、検索の目的によっては、その一致区間の正確な位置と長さに基づいて分類まで行いたいこともある。その場合、どの区間がどの区間と一致するかを明確に検索結果として出力できなくてはならない。また、図中のNo.2とNo.3にあるように、重なりがある場合には、その包含関係を考慮した上で示せなければならない。本発明の手法では、この課題についても基本的な照合原理を変更することなく高速処理が可能である。本手法における照合処理では、対象映像からフレームが入力されて特徴量が得られると、問合せ用映像中から、それと同じ特徴量を持つ一群のフレームが候補に入れられると述べた。このとき、同時に候補に入ったフレームを先頭とし、検出閾値以上の長さに達した一致区間の一群は、互いに同一の映像である。図7の例では、②の区間は3つの問合せ用映像のいずれにも存在し、各問合せ用映像における区間先頭フレームはすべて、対象映像から②の先頭に当たるフレームが入力された時点で同時に候補に入る。他にも同時に候補に入るフレームがある可能性があるが、それらは検出閾値以上の長さに達する前に候補から外れることになる。そして、②の区間の末尾まで到達し、その次のフレームの照合を行った時点で、No.1とNo.3の問合せ用映像における一致区間が候補から外される。対象映像にとっては、まだNo.2の間では一致しつづけているが、ここでひとまず②の区間を確定し、問合せ用映像No.1~3で②が見つかったことを検索結果として出力する。しかし、問合せ用映像No.2は、区間②が終わっても、次のフレームもまだ対象映像と一致しているので候補として残り続け、最終的に区間③が確定する。また、①のように、②より手前に区間があっても同様に一致区間が検出され確定する。このように、本発明の手法によれば、候補に入るときと外れるときに簡単なチェックを行うだけ

で、フレーム入力ごとの照合の処理量は少ないまま、微妙に前後が違っている様々なバリエーションのシーンをそれぞれ区別して検出することができる。

【0022】以上の説明では、問合せ用映像をあらかじめ用意しておき、その上で対象映像から検索を行うケースについて説明したが、本手法は問合せ用映像が対象映像そのものであっても適用できる。図8にその概念図を示す。対象映像を入力し、それをすべて記憶しておき、それをあたかも上述の問合せ用映像であるかのように扱う。これは、図9のブロック図によって実現することができる。ほとんど図2のブロック図と変わらないが、問合せ用映像と対象映像とが同一なため、フレーム特徴量の抽出までが共用でき、フレーム特徴量108が記憶用と照合用にそれぞれ分配される。この機構により、対象映像から入力された最新の映像部分①が、過去に入力されたどの部分に現れているかを入力と同時に検出することができる。また、過去に複数回現れたシーンであれば、上述の照合原理から、それらすべてが同時に検出されるので、検出された同一シーンごとにまとめて分類整理する、いわば映像の自己組織化がリアルタイムで自動的に成される。たとえば、数週間分のテレビ番組を録画できる装置に、本発明を数週間分のテレビ番組をすべて記憶できるだけのメモリ容量を持たせて適用すれば、一般に番組のオープニングには毎度同じ映像が流れるので、その映像を検出して前後の映像をまとめることで番組の分類整理が録画と同時にリアルタイムで行える。同じシーンが複数あることがわかれば、映像そのものは1つだけ残しておいて、残りについてはポインタだけ残して消してしまえるので、録画するときのメディアの利用効率を向上させることができる。また、コマーシャルも繰り返し流される映像の一つであるが、録画した番組を再生する場合に、必要に応じて自動でスキップさせることもできる。このとき、コマーシャルの特徴として、長さが15秒もしくは30秒ちょうどであることを利用すれば、コマーシャルかどうかの判定性能が向上する。

【0023】以上説明した中で、図9で示したブロック図を実現する処理について、より具体的にフローチャートで表現すると、図10のようになる。図2のブロック図を実現する処理についても、図10より自明となる。また、上記では説明を簡単にするため、問い合わせ用映像の特徴量を一度ラン単位からフレーム単位に戻して照合したが、ここではより実用に近い仕様となるよう、ラン単位のままで照合する方法を示す。

【0024】まず処理200で装置や各種変数の初期化を行う。変数 $mc$ と $mn$ に0が入る。次に、対象映像からフレーム画像を1枚入力し(202)、そのフレーム画像から特徴量 $F$ を抽出する(204)。特徴量 $F$ は、フレーム画像中に存在する全画素の色の平均を用いる。各画素の色はRGB3つのコンポーネントから表現されているが、各コンポーネントの値についてそれぞれ全画面分

の平均をとり、3つの値の組(Ra, Ga, Ba)を得、この組をもって特徴量Fとなす。もし、初めてのフレーム入力ならば、図11に示した特徴量テーブル構造体300を新規に作成し、最初の区間(区間番号1)の特徴量として302にFを書き込む。また、このときのフレーム番号も対にして一緒に304に書き込む。このようにして作成された特徴量テーブルが、今後既に述べた問合せ用映像として機能することになる。このとき、特徴量テーブル300に記憶されている区間の最大値を示す変数mcを1つ増やし、そのまま202に戻る。一方、2回目以降のフレーム入力ならば、処理206を行う。206では、特徴量テーブルに記憶された最新の区間(区間番号mc-1の区間)の特徴量FCと今回の特徴量Fとを比較し、その差異が閾値CTH以下かどうかを判定する。ここで、特徴量は上述のように3つの値の組であるが、3つの値の差がすべて閾値CTH以下であるときのみ、差異が閾値CTH以下と表現することにする。もし、差異が閾値CTH以下であれば、今回入力されたフレームは、直前までのフレームと同じ区間にまとめることができると判断され、処理208に進む。208ではループカウンタiを0にリセットする。iは226で1ずつ増やされ、iがmmより大きくなるまで210~224の処理を繰り返す。ここでmmは、これまでに入力した全映像(特徴量テーブル300として記憶)の中で、現在新たに入力されつつある映像と同じ部分である可能性があるとして継続審査の段階にある候補の数である。すべての候補について、それぞれ審査段階を表わす状態変数を格納する構造体500が作られており、図12に示すような候補リスト構造体400によって管理されている。400には、候補構造体500へのポインタが格納され、追加や削除が実行中に動的に行われる。図13は、候補構造体500の構造を示しており、候補として登録されたときの区間番号が照合開始区間番号502として、その区間からスタートして現在照合の対象となっている区間の番号が照合対象区間番号504として格納されている。また、一致フレーム数カウンタ506は、候補に入ってから一致しつづけた回数、すなわち一致区間の長さである。照合開始フレームオフセット508は、ラン単位で比較しながらフレーム精度での位置決めを行うために必要な変数であり、後述する。そして、同時照合開始候補へのポインタ510は、同時に登録された一群の候補どうしを接続リスト形式で結んでおり、510を参照していくことで同時に登録された候補を次々に辿ることができる。さて、処理210では、候補i(mm個ある候補の中のi番目の候補という意味で表記)が、現在照合対象となっている区間の末尾まで照合し終わったのかどうかを調べる。これは照合開始区間番号502で示される区間のフレーム番号に、一致フレーム数カウンタ506を足したフレーム番号が、現在照合対象となっている区間の次の区間のフレーム番号に達していれば、末尾に

達しているとわかる。もし、達していなければ、候補iの一致フレーム数カウンタを1増やして(216)、処理226に進む。達していれば、現在照合対象となっている区間に後続する区間の特徴量を参照し、それとFとの差異が閾値STH以下かどうかを調べる(212)。もし、差異が閾値STH以下ならば、照合対象の区間をその後続区間に変更して照合を継続する(214)。これによって、区間が変わる場所が入力映像と異なっている場合でも安定して照合ができる。これは、映像入力時のノイズや機器の特性等によって映像信号が変動することがあるため、区間の変化点が、同じ映像を入力していても必ずしも同一になるとは限らないために必要な処理である。また、区間の変わり目を決定する閾値CTHと異なる閾値STHをここで用いているのも、同様に映像の変動を吸収し、安定した照合を行うためである。一方、処理212で、差異が閾値STHより大きければ、現在照合対象となっている区間の特徴量と、今回の特徴量Fとの差異が閾値STH以下かどうかを調べる(218)。もし、差異が閾値STH以下ならば、何もせずに処理226に進む。これは、候補に入るのがフレーム単位でなく区間単位であり、必ずしも区間の先頭から一致するとは限らないために、現在照合対象の区間と同じ特徴量の入力映像が得られている間は、とりあえず位置合わせしながら待ってみるのである。もし、差異が閾値STHより大きければ、もう一致しなくなったみなす。このとき、候補iの一致フレーム数カウンタが閾値FTH以上ならば(220)、検索シーンとして候補iを出力する(222)。そして、候補iを候補リストから削除し(224)、処理226に進む。

【0025】さて、処理206において、差異が閾値CTHより大きければ、今回入力されたフレームは、直前までのフレームと同じ区間にまとめることができないと判断され、新規区間を特徴量テーブル300に追加する(228)。このとき、mcを1つ増やし、FCにはFを代入しておく。230ではループカウンタiを0にリセットする。iは248で1ずつ増やされ、iがmmより大きくなるまで232~246の処理を繰り返す。処理232では、候補iが、現在照合対象となっている区間の末尾まで照合し終わったのかどうかを調べる。これは処理210と同様の方法で求められる。もし達していれば、照合対象の区間をその後続する区間に変更し(234)、そうでなければ何もしない。次に現在照合対象となっている区間の特徴量と、最新の特徴量Fとの差異が閾値STH以下かどうかを調べる(236)。もし、差異が閾値STH以下ならば、候補iの一致フレーム数カウンタを1増やし(238)、処理248に進む。もし、差異が閾値STHより大きければ、現在照合対象となっている区間の後続区間について直後の1つだけでなく順番に調べていき、今回の特徴量Fと同じ特徴量を持つ区間があるかを調べる(240)。もしあれば、その後続区間を

照合対象の区間に変更し、また、その区間のフレーム番号と当初の照合しようとしていたフレーム番号との差を照合開始フレームオフセット508に代入してから、処理248に進む。これもまた必ずしも区間の先頭から一致するとは限らないため、このオフセットを利用することでフレーム精度の位置合わせができるようになっている。このとき、オフセットの大きさが、候補に入れられたときの区間の長さを超えていたら、一致する後続区間が見つからなかったと同じ扱いで処理242に進む。さもなければ、最初に候補に入ったときの区間の後方の区間から照合開始したと等価であり、そのときは、その後方の区間から照合開始したものが順調に一致を続けているはずで処理が重複することになる。一致する後続区間が見つからなかった場合、候補iの一致フレーム数カウンタが閾値FTH以上ならば(242)、検索シーンとして候補iを出力する(244)。そして、候補iを候補リストから削除し(246)、処理248に進む。全部の候補について処理が終わったら、今回入力されたフレーム画像の特徴量と同じ特徴量を持つ区間を、特徴量テーブルに記憶された区間の中からすべて探し出し、それらの区間を照合開始区間とする候補構造体を作成して候補リストに追加する(250~256)。

【0026】以上の処理の中の222、244において、見つかったシーンの情報をそのまま出力するほかに、図14に示した形式で出力することもできる。600は検索結果テーブルで、見つかったシーンを同一シーンごとにまとめてグループを作り、各グループのエントリを管理する。同一シーンのグループは、先に図7で説明したようにして得られる。見つかったシーンの一つ一つは検索区間構造体700で表現され、同一シーンは相互にポインタを持ち合う接続リスト形式で一つのグループを表現する。接続リストを形成する同一シーンへのポインタは704に格納され、また、それぞれの区間の先頭フレーム番号が702に格納される。そして、グループを表わす接続リストの先頭となった検索区間構造体へのポインタが、そのグループのエントリとして602に格納される。同一グループでは、グループ内のすべてのシーンの区間長は同じなので、エントリと対にして604に格納する。

【0027】以上の処理を繰り返すことで、過去に一度でも現れたことのあるシーンは再び現れた瞬間に検出され、その区間の先頭と長さがフレーム精度で位置決めされる。区間の先頭は、候補構造体の照合開始区間番号で示される区間のフレーム番号に、同じく照合開始フレームオフセットを加えたフレームであり、長さは、一致フレーム数カウンタの値そのものである。あとは先に図7を用いて説明したように、同じ区間ごとにまとめることで自動自己組織化ができる。但し、静止画が長く続くようなシーンの場合、フレーム1枚あたりの特徴量を減らしている本手法では特徴量の特徴的な時間変化が得られ

ず、他の静止画シーンと間違っ一致する確率が高くなるという問題もある。この場合については、たとえばフレーム画像1枚あたりの特徴量を増やすことで対応できるのはいうまでもない。また、同様に特徴量に変化が少ないシーンの場合には、数フレーム程度のずれがあっても一致しうる。このような場合、複数の区間が同じ範囲に重なり合って検出される。この典型的な例として、入力されたばかりの映像が、同じカット(映像を構成する単位の一つ。1台のカメラで途切れなく撮影されたひとまとまりの映像区間)内の少しだけ前のほうの区間と一致する場合がある。映像の冗長性から、同じカット内の各フレームは画像的に良く似ているためである。これには、公知のカットの変わり目の検出手法を導入し、同じカット内では一致とみなさないといった処理を行うことで問題を回避できる。

【0028】図15は、本発明、特に図8に示した手法を利用した、次世代ビデオレコーダシステムの一実施例を示す概念図である。テレビ番組などの映像を録画すると同時に、本発明の機能も実行する。録画する映像の各フレームにはフレーム番号等のアドレス情報を付与し、そのアドレス情報を、本発明で生成する特徴量テーブル300のフレーム番号304として用い、映像データと特徴量テーブルとの1対1の同期をとる。録画終了時には、本発明で用いる特徴量テーブルや各種変数は不揮発性の記録装置に記録され、次回録画開始時に読み出されて再開できるようにする。これにより、映像を新しく入力すると同時に、映像アーカイブの中にすでに記録された映像とリアルタイムで照合し、同一シーンどうしを自動的に関連づけることができる。そして、例えば、入力された映像と主題歌の部分が照合する番組がすでに格納されていれば、それらは連続番組であり、同じ分類としてまとめて整理することが自動的にできる。もし、連続番組の初回を見たときに、その連続番組全般の共通属性として情報を付与したならば、入力されたばかりの映像にも、その情報を直ちに共有させることができる。また、繰り返し現れるコマーシャルも同様にして検出でき、コマーシャルをスキップすることができると先に述べた。しかし、録画格納されている映像に存在するコマーシャルをもとにしただけでは、限られた数のコマーシャルしか検出できない。そこで、映像を録画しないときも24時間映像を調べるようにし、繰り返しシーンからコマーシャル部分を検出し、そのコマーシャル部分の映像については、映像の録画はしないまでも特徴量テーブルだけは作成し記録しておくようにする。これにより、映像の容量を抑えたまま、より多くのコマーシャルが検出できるようになり、また、コマーシャルのスキップがより確実に行えるようになる。このように、本発明を次世代ビデオレコーダシステムに搭載することで、録画した番組を自動分類整理したり、コマーシャルを自動スキップしたり、といったことが簡単にでき、非常に使い勝

手がよくなる。ここでは放送映像を対象とできることを強調したが、ファイル等に格納された映像が対象であっても構わないことはいうまでもない。

【0029】図16はユーザとの対話に利用する表示画面の一実施例である。映像のフレーム画像はコンピュータのディスプレイ上のモニタウインドウ50に再生表示される。同じ画面上に表示されるウインドウには、50のほか、映像中の代表的なフレーム画像を一覧にして表示するウインドウ52、映像やシーンの属性を入力するテキストウインドウ55、検索結果を表示するウインドウ54などがある。検索結果が52に表示される場合もある。これらのウインドウは、ポインティングデバイス3の一つであるマウスを使って自由に動かせるカーソル53を操作して画面上の任意の位置に移動できる。テキストの入力にはキーボード4などを用いる。ウインドウ52に表示される代表フレームは、例えば、映像をカット単位に分割した際の各カットの先頭フレームとする。ボタン51は、映像の再生状態を制御するためのボタンで、マウスでクリックすると、映像の再生や早送り・巻き戻しといった制御ができる。また、再生するシーンについては、ウインドウ52に一覧表示されている代表フレーム画像をクリックすることによって連動的に選択を行うこともできる。そのとき、再生するビデオは、コンピュータに接続されたビデオ再生装置5が出力する映像を用いることもできるし、外部情報記憶装置に登録されたデジタル化された映像でもよい。ビデオ再生装置5の場合には、シーンの先頭のフレーム番号をビデオ再生装置に送り、そのフレーム番号に対応するシーンから再生を開始する。そして、シーン末尾のフレーム番号に到達した時点で再生を中断する命令をビデオ再生装置5に送る。デジタル化された映像の場合も基本的に同様だが、デジタル映像データを読み取った後、計算機向けの描画データに変換してグラフィックの一種として表示する。1フレーム分の表示処理が終了すると、次のフレームの表示処理を連続して行い、これにより動画像の表示とする。表示処理に要する時間に合わせ、一定時間あたりに表示するフレーム画像の枚数を調節し、映像が早送り気味になったりスロー気味になったりしないようにする。またモニタウインドウ50には放送受信機7からの映像の表示も行える。

【0030】ユーザが図16の画面を用いて映像検索を行う操作手順を以下のべる。まず問合せ用映像の指定を行う。最も単純な方法は、51の操作ボタンを用いて早送りや巻き戻しを行い、モニタウインドウ50に映される映像を確認しながら任意のシーンを見つける方法である。また、ウインドウ52に並んだ代表フレームの一覧は、本でいうところの目次や索引に相当するもので、それを参照することで、より手早く所望のシーンを見つけることができる。指定する際には、そのシーンの範囲を正確に指定する必要はなく、そのシーンに含まれる任意

のフレームを指定すればよい。これはモニタウインドウ50に表示中のフレームをマウスでクリックして指定するのでもよいし、52の代表フレーム一覧の中に、問合せ用映像に含まれるフレーム画像が表示されているならば、それをマウスクリックするのでもよい。次にテキストウインドウ55で、選択したシーンやその映像全体のタイトルや人名等の属性情報をキーボードから入力し登録する。尚、登録数は任意であり、また、属性情報を今後再利用する必要がない場合には、属性情報の登録は一切行う必要がない。最後にユーザは検索開始要求を出す。これは55のOKボタンをクリックすることで行う。これにより、システムは検索処理を開始する。システムは、指定されたフレームをちょうど中間とする一定の長さの区間を仮想的に作り、その区間を問合せ用映像として本発明の検索手法にかける。対象映像はビデオ再生装置から新たに入力してもよいし、すでにデータベースとして登録され特徴量テーブルが作成されている映像ならば、その特徴量テーブルに対して照合処理を行う。このとき、得られた検索結果の区間の中に、最初に指定したフレームが含まれていれば、それが検索結果となる。さらに、部分一致なのか、全区間一致なのかを調べて、全区間一致しているときには、区間を前方および後方に広げて一致区間を正確に求めることもできる。これは、部分的に一致する区間も高速に探し出せる本発明の手法のメリットを活かした検索方法となっている。

【0031】検索結果はウインドウ54に表示される。表示内容は属性情報、時間情報等である。あるいは検索結果を、図17に示した形でグラフィカルに表示することも可能である。図17は、ウインドウ52の拡大図になっており、800が各代表フレームのアイコン画像である。横棒806がアイコン画像の下に添えられることで、そのアイコン画像に対応するシーンに、検索結果が存在していることがわかるようになっている。検索結果が、アイコン画像の複数のシーンにまたがれば、棒もその分だけ長くなる。また、棒は色もしくはハッチングパターンで分類されており、同じシーンの検索で見つかった複数のシーンはいずれも同じ色で、一方、あるシーンの検索結果と別のシーンの検索結果とは違う色で表示される。代表フレーム一覧は、前述のように映像の目次や索引として用いることができ、問合せ用映像を見つけるのに非常に有用であるが、代表フレームは映像に含まれるすべての画像ではなく、かといって、すべての画像を一覧にすれば、それらの中から所望のものを見つけるのが困難になる、というジレンマがある。そのため、映像を解析することで、代表フレームが表わすシーンの典型的な特徴を抽出し、例えば、各アイコン画像800に、特徴を表現する情報802や時間情報804を合わせて表示することで、代表フレームの画像にない部分の映像も探し出せるようにすることが考えられる。このようなシーンの特徴を表現する情報には、人物の有無やカ

メラワーク（ズームやパン、チルトなど）、特殊効果（フェードイン／アウト、デゾルブ、ワイプなど）の有無、字幕の有無、などがある。画像の検出を行う画像認識手法には、先に本発明者らが特許出願した特願平7-210409号（H7.8.18出願）を利用すればよい。本発明の手法を適用すれば、また別のアプローチから、代表フレーム一覧のジレンマの解消に役立てることができる。代表フレーム一覧には、繰り返しシーンであれば、それらのシーンの全部ではなくても、その幾つかが一覧の中に含まれていることがある。例えば、図18において、カーソル53で、繰り返しシーンの一枚をクリックし検索を行えば、そのシーンと同じ映像部分を持つシーンをすべて見つけて、ユーザに提示するようにする。検索結果の提示は、たとえば、アイコン画像808に重畳表示した星印810のように、検索された区間を含むシーンのアイコン画像を強調する形で行う。このとき、表示するアイコン画像自体も、検索された区間の中のフレーム画像で置き換えると、さらにわかりやすくなる。これによって、代表フレーム一覧の中に、見つけたいシーンと同じシーンの画像が一つでもあれば、それを頼りに所望のシーンを見つけてことが可能になり、代表フレーム一覧の利便性が高まる。同様の手法は、モニタウインドウ50で表示中の映像についても適用でき、表示中のフレームをクリックするなどして指定して、そのフレームを含むシーンと同じシーンを検索し、見つかったシーンの一つにジャンプさせるといったこともできる。このようなことを実現するためには、これまでリンクノードの設定などの面倒な準備が予め必要であったが、本発明の手法を用いれば、非常に高速に検索ができるので、必要な都度検索を行えばよく、前もって準備しておく必要はない。

【0032】尚、図9のブロック図で表わされる自己組織化処理の実行にあたっては、ユーザは検索のための特別な処理を何ら行う必要はなく、ただ映像を入力しさえすれば計算機が自動的に処理を実行する。

【0033】また、以上の説明では、映像の画像特徴に基づき検索を行う方法について述べたが、音声特徴を用いてもよく、また、映像に限らず、逐次的に扱うことのできるメディアであれば、いずれにでも本検索手法を適用できることは言うまでもない。

【0034】図19に本発明の画像検索技術をビデオカメラに適用した例を示す。処理入力手段1960に設けられた電源スイッチ1961により電源が投入され、録画ボタン1962により録画が指示されると、音声・画像入力手段1910は、マイク1911から音声信号を、カメラ1912から画像信号を入力する処理を行う。音声・画像入力手段における処理には、入力された音声・画像信号に対してA/D変換処理、圧縮処理が含まれる。特徴抽出手段1970は入力された画像信号からフレーム単位の特徴を抽出する。処理の内容は、図2、図9の106と同じである。抽出された特徴は特徴量テーブルとしてメモリ1940に記憶され

る。メモリ1940は内蔵された半導体メモリ、着脱式のメモリカードを用いる。メモリ1940には入力された音声・画像信号が保持され、再生ボタン1963からの再生指示により、メモリ1940から読み出され、音声・画像出力処理手段において、信号圧縮に対する伸長処理や、D/A変換処理が施され、画像は表示画面1921に出力され、音声はスピーカ1922から出力される。制御手段1930はこのビデオカメラの信号処理全般を管理し制御する。入力された画像はフレーム毎にその特徴が抽出されメモリに蓄積されるが、制御手段1930では入力画像の特徴量をメモリ1940中に保持されている過去のフレームの特徴量と照合する。照合処理は、図2、図9の特徴量照合処理130と同じようにすればよい。照合した結果、特徴量が類似するシーンの区間が検索結果テーブル（図2、図9の128）と同様の形式でメモリ1940に保持される。1950はビデオカメラを駆動するための電源を供給する端子であり、バッテリーを装着するようにしてもよい。画像検索メニューボタン1964は、録画された動画像を例えば図16、図17、図18のように表示画面上1921に表示された画面で、ボタン1964を複数回押すことにより、シーンの並び替えや削除等の簡単な編集処理や、所望のシーンを指示して類似のシーンにを検索して再生させるといった処理を指示する。シーンの区分けに用いられる動画像の変化点検出に関する技術は先に発明者らが特許出願した特願平成7-32027号（H7.2.21出願）を参照すればよい。シーンの検索は、図2、図9で行った画像の特徴量の照合処理を用いて行う。このようなビデオカメラは、特徴量の照合処理の条件をゆるめに調整する必要がある。テレビ番組とは違い、一般にユーザがビデオカメラで映像を撮影するときには、全く同じ映像を撮影することはほとんどないからである。従って、同じ様な風景や、同じ服装の人が同程度の大きさで写ったときには、類似シーンとして検索されるように照合条件を設定する。録画と同時に撮影された画像が解析され、シーン毎のグループ分けと類似シーン間でのインデックス付けが完了しており、撮影直後に録画画像の編集が可能になり、ユーザの使い勝手が向上する。

【0035】

【発明の効果】上記の方法により、本発明では、ほとんど同じ特徴量が続く冗長な区間をひとつにまとめた単位に照合を行うので、照合を毎フレームについて行う必要がなくなり、大幅に計算量が削減できる同時に、照合にあたっては、擬似的にフレーム単位の特徴列間で比較する形になるため、フレーム精度で同一映像区間の特定を行える特徴がある。また、1回のフレーム入力があるたびに、その1枚のフレームに関してのみの照合を行うため、1フレーム入力あたりの処理量が小さくなり、放送映像をはじめとしたリアルタイム性が要求される映像の処理に好適である。また、同時に検出された複数の映像部分は、全く同じ映像であるので、それらを1つの組と

して記憶格納しておくことにより、1つの部分映像を探したい要求があった場合には、組になった他の部分映像を提示することで検索は完了し、極めて高速な応答が期待できる。

【図面の簡単な説明】

【図1】本発明の一実施例を実行するシステムのブロック図である。

【図2】本発明の一実施例を実行する処理のブロック図である。

【図3】本発明の一実施例の特徴量抽出方法を示す概略図である。

【図4】本発明の一実施例の特徴量照合方法を示す概略図である。

【図5】本発明の一実施例の特徴量照合の流れの一例を示す図である。

【図6】従来の照合方法の一例を示す概略図である。

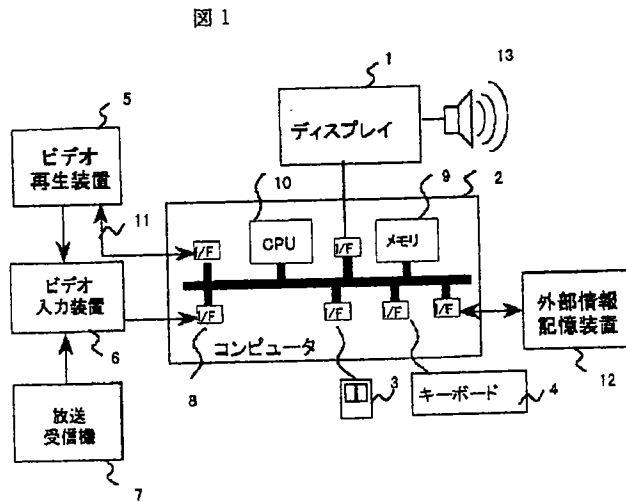
【図7】本発明の一実施例の照合方法を説明するための概略図である。

【図8】本発明の一実施例の照合方法を説明するための概略図である。

【図9】本発明の一実施例を実行する処理のブロック図である。

【図10】本発明の一実施例のフローチャートである。\*

【図1】



\*【図11】本発明の一実施例で使われる特徴量テーブル構造体を示す図である。

【図12】本発明の一実施例で使われる候補リスト構造体を示す図である。

【図13】本発明の一実施例で使われる候補構造体を示す図である。

【図14】本発明の一実施例で使われる検索結果テーブルと検索区間構造体を示す図である。

【図15】本発明の一実施例を応用したビデオレコーダシステムの概略図である。

【図16】本発明により自己組織化された映像の映像検索時の表示画面例を示す図である。

【図17】本発明により自己組織化された映像の映像検索時の表示画面例を示す図である。

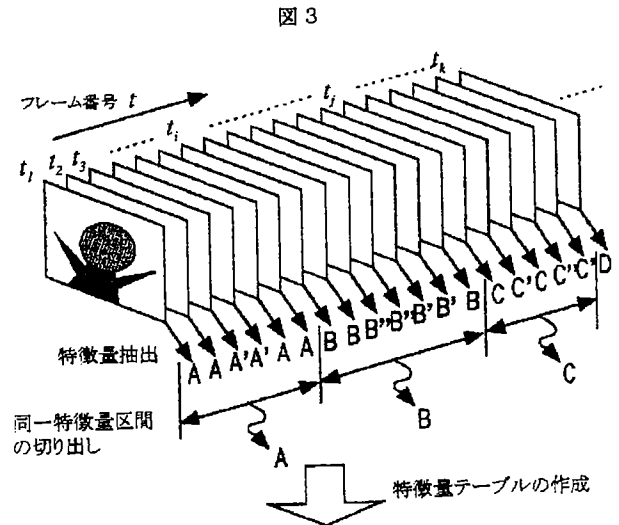
【図18】本発明により自己組織化された映像の映像検索時の表示画面例を示す図である。

【図19】本発明をビデオカメラに適用したときの構成概略図である。

【符号の説明】

1…ディスプレイ、2…コンピュータ、3…ポインティングデバイス、4…キーボード、5…ビデオ再生装置、6…ビデオ入力装置、7…放送受信機、12…外部情報記憶装置。

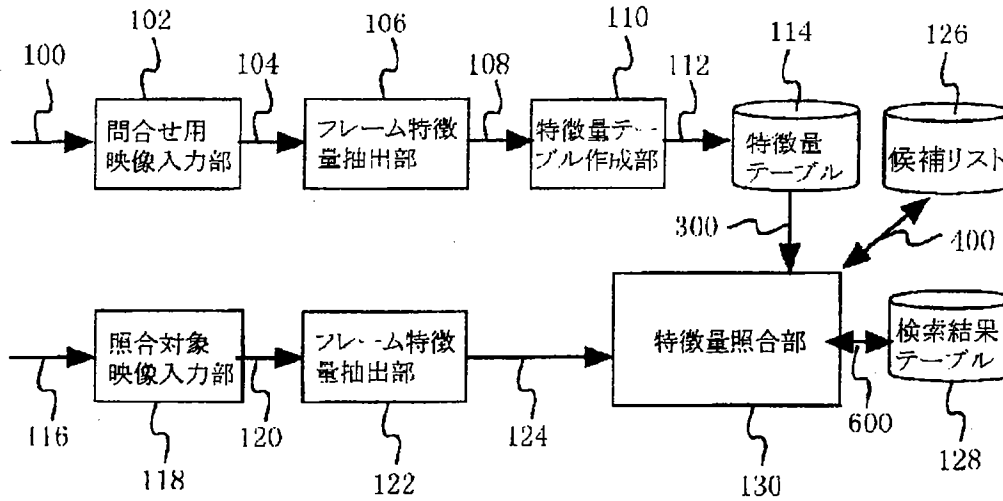
【図3】



区間番号	特徴量	フレーム番号
1	A	$t_1$
2	B	$t_1$
3	C	$t_1$
4	D	$t_k$
⋮	⋮	⋮

【図2】

図 2

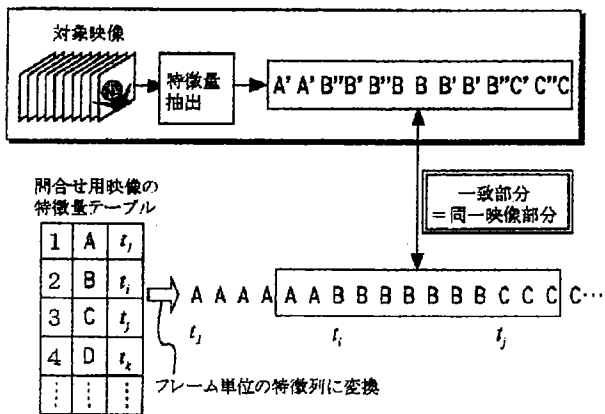


【図4】

【図5】

図 4

図 5



【図11】

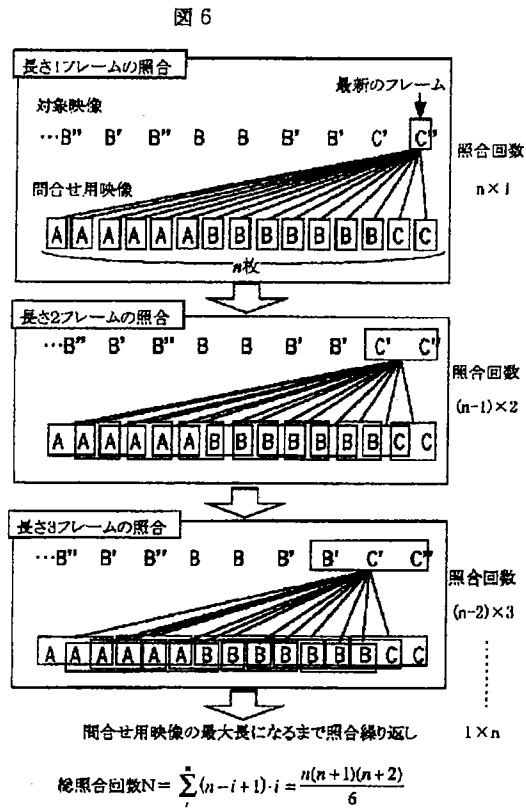
図 1 1

特徴量テーブル構造体

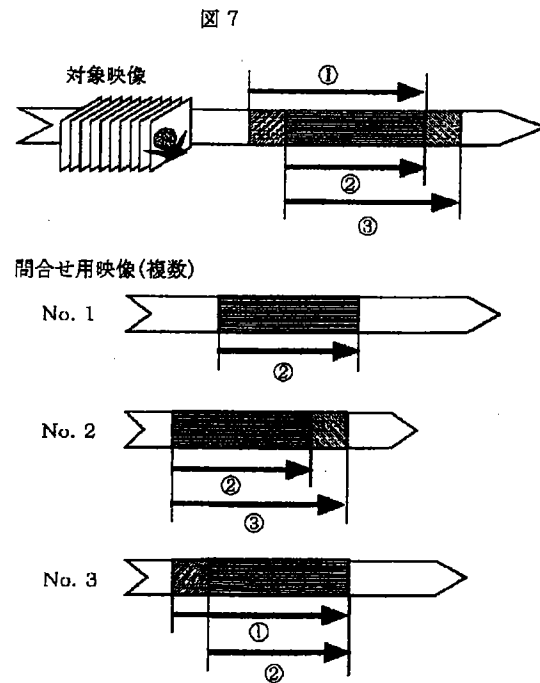
区間番号	特徴量	フレーム番号
1	A	$t_i$
2	B	$t_j$
3	C	$t_j$
4	D	$t_k$
...	...	...

フレーム	候補フレーム																候補数
(1)																	0
(2)																	0
(3)	①	②	③	④													4
(4)	1	1	1	x	①	②	③	④									7
(5)	x	x	1		x	x	x	1	⑤	⑥							4
(6)		1				1	1	x	⑤	⑥							5
(7)		1				1	1	x	1	⑦	⑧						5
(8)		1				1	1		1	1	x	⑦	⑧				8
(9)		x				x	x		x	x		x	x				0

【図6】

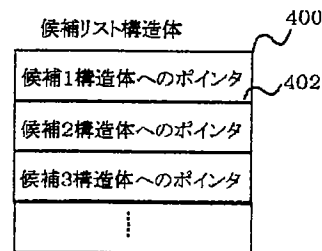


【図7】



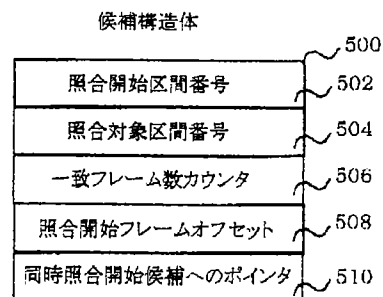
【図12】

図12



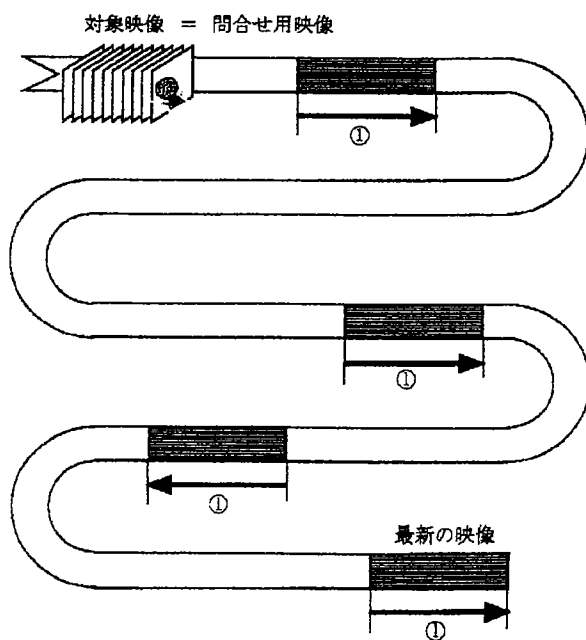
【図13】

図13



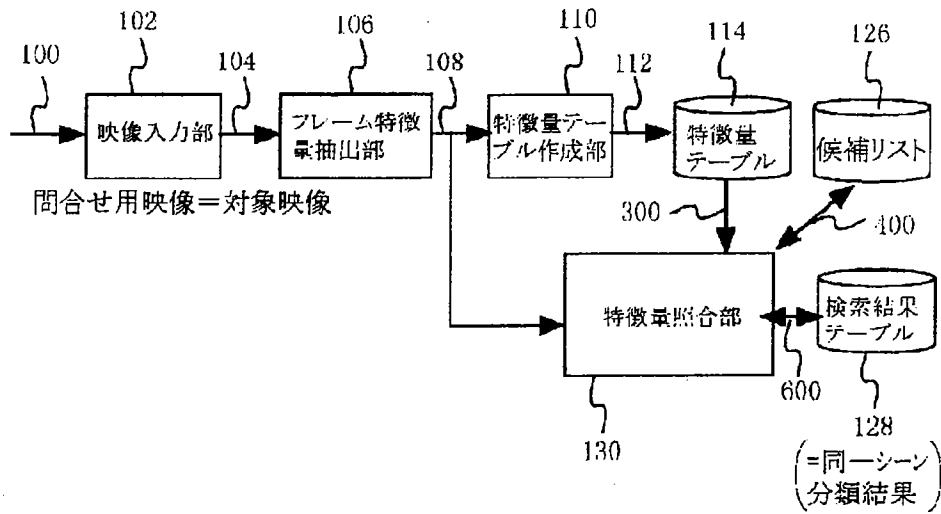
【図8】

図8



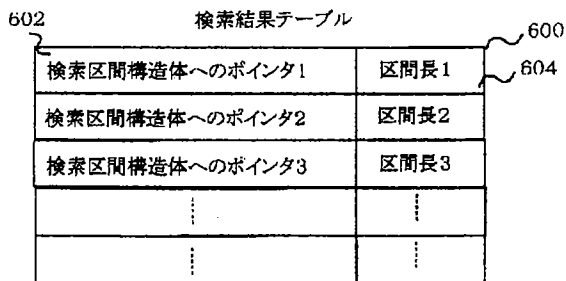
【图9】

图 9



【圖 14】

图 14



【図15】

15

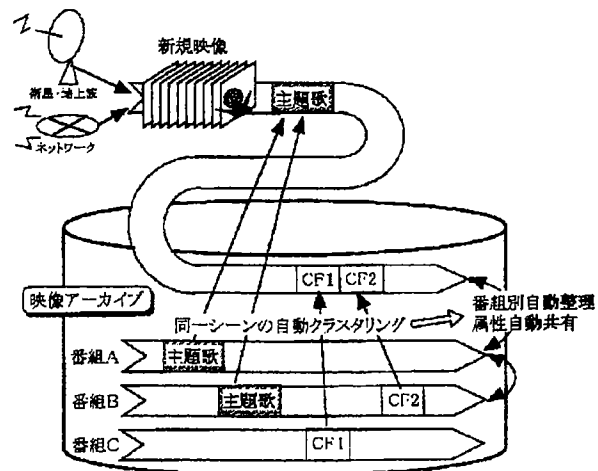
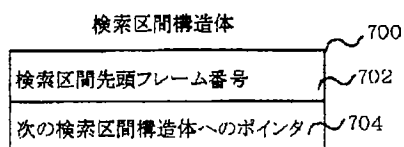
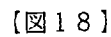
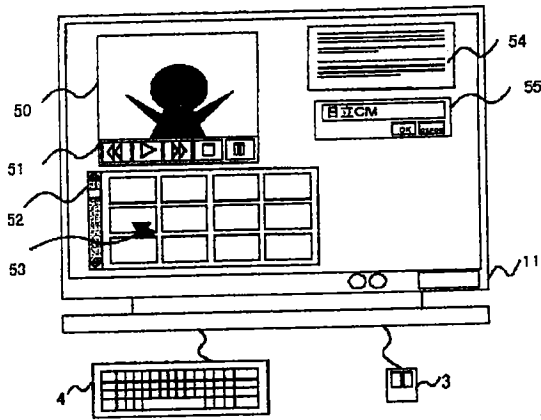


图 10



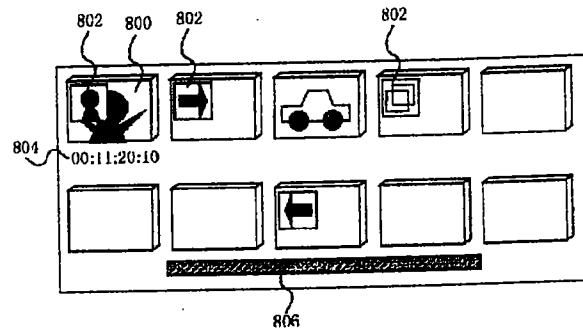
【図16】

図16



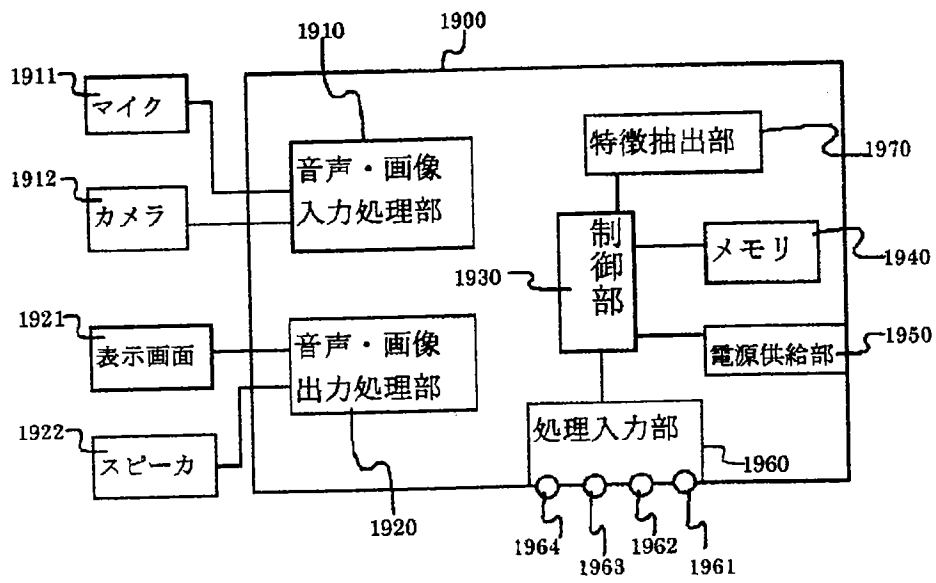
【図17】

図17



【図19】

図19



【公報種別】特許法第17条の2の規定による補正の掲載  
【部門区分】第6部門第3区分  
【発行日】平成15年1月17日(2003.1.17)

【公開番号】特開平10-320400  
【公開日】平成10年12月4日(1998.12.4)  
【年通号数】公開特許公報10-3204  
【出願番号】特願平9-126704  
【国際特許分類第7版】  
G06F 17/30  
【FI】  
G06F 15/40 370 D

【手続補正書】  
【提出日】平成14年10月10日(2002.10.10)  
【手続補正1】  
【補正対象書類名】明細書  
【補正対象項目名】発明の名称  
【補正方法】変更  
【補正内容】  
【発明の名称】 動画像の記述方法  
【手続補正2】  
【補正対象書類名】明細書  
【補正対象項目名】特許請求の範囲  
【補正方法】変更  
【補正内容】  
【特許請求の範囲】  
【請求項1】動画像から所定区間毎に特徴量を抽出するステップと、前記抽出された特徴量を比較して、前記特徴量の変動が許容値内にある連続する区間の代表特徴量を得るステップと、前記代表特徴量を該代表特徴量によって代表される区間長を示す情報とともに記憶するステップとを有することを特徴とする動画像の記述方法。  
【請求項2】動画像から所定区間毎に特徴量を抽出するステップと、前記抽出された特徴量を比較して、前記特徴量の変動が許容値内にある連続する区間の代表特徴量を得るステップと、前記代表特徴量を該代表特徴量によって代表される区間の時間に関する情報とともに記憶す

るステップとを有することを特徴とする動画像の記述方法。

【請求項3】動画像から所定区間毎に特徴量を抽出するステップと、第1の前記特徴量を代表特徴量として、及び、前記第1の特徴量の間隔を第1の区間として設定するステップと、対象となる前記抽出した特徴量を前記代表特徴量と比較するステップと、前記比較の結果が、許容量以内の場合には前記第1の区間に前記抽出した特徴量の区間を加え、許容量を越える場合には、前記抽出した特徴量を別の代表特徴量と、及び、前記抽出した特徴量の区間を他の区間として設定するステップと、前記代表特徴量によって代表される区間を示す区間の時間に関する情報とともに前記代表特徴量を記憶するステップとを有することを特徴とする動画像の記述方法。

【請求項4】前記特徴量として色を用いることを特徴とする請求項1乃至3の何れかに記載の動画像の記述方法。

【請求項5】前記抽出するステップは、前記動画像を構成するフレーム画像毎に特徴量を抽出することを特徴とする請求項1乃至4の何れかに記載の動画像の記述方法。

【請求項6】前記区間の時間に関する情報は、フレーム番号もしくはフレーム記録時刻であることを特徴とする請求項1乃至5の何れかに記載の動画像の記述方法。